Statistics Canada | Statistique Canada

# Social and Economic Studies Division

# Division des études sociales et économiques

Canada

E-620F-02

# ANALYSE DE LA MOBILITÉ LINGUISTIQUE: INDICES, OBSERVATIONS ET MODÈLES

par

Réjean Lachapelle*

Document de recherche n° 12a, mai 1985
Division des études sociales et économiques
Statistique Canada

Article soumis aux Cahiers québécois de démographie.

Décembre 1984

* Nous avons tenu compte dans toute la mesure du possible des commentaires de Leroy O. Stone. Nous remercions aussi Luc Albert et Luc St-Amour pour leur soutien technique ainsi qu'Audrey Sirois qui a assuré la frappe du texte. L'auteur, et non Statistique Canada, assume la responsabilité pleine et entière des idées exprimées dans le présent document.

**Sommaire**

Deux méthodes d'analyse de la mobilité linguistique sont présentées. La première repose sur la distinction entre d'une part la mobilité structurelle, laquelle provoque un changement de la composition de destination (ou par langue parlée au foyer) par rapport à la composition d'origine (ou par langue maternelle), et d'autre part la mobilité d'échange, qui mesure l'intensité des transferts équilibrés entre chaque couple de langues. Les deux séries de paramètres qui y correspondent, complétées par la composition selon la langue maternelle, épuisent la totalité de l'information que renferme le tableau fondamental dans lequel se trouve consignée la distribution conjointe de la population selon la langue maternelle et la langue parlée à la maison. Tous les autres indices peuvent donc être définis à l'aide de ces coefficients. Cela permet de résoudre les paradoxes que les études sur la mobilité linguistique ont mis en évidence. La seconde méthode d'analyse est plus classique puisqu'elle repose sur la théorie des chaînes de Markov stationnaires. Elle consiste à transformer les taux de mobilité linguistique en une composition d'équilibre. Celle-ci est ensuite comparée aux compositions observées. Sauf pour certaines questions bien spéciales, cette méthode procure des résultats moins intéressants que la méthode précédente, laquelle est mieux adaptée aux préoccupations qui sous-tendent les travaux sur la mobilité linguistique.

## Introduction

Depuis la parution des données du recensement de 1971 touchant la langue maternelle(1) et la langue au foyer (ou langue parlée à la maison), de nombreux travaux furent consacrés à l'étude de la mobilité linguistique et aux événements qu'elle engendre, les transferts linguistiques (voir notamment: Castonguay, 1983a; De Vries et Vallee, 1980; Gryz, 1980; Lachapelle et Henripin, 1980). Plusieurs stratégies d'analyse furent distinguées et mises en oeuvre et l'on a proposé et utilisé de nombreux indices pour mettre en évidence les différentes facettes de ce phénomène. La diversité des stratégies et des méthodes a certes permis d'asseoir certains résultats sur des bases solides, mais elle a aussi conduit à des conlusions en apparence contradictoires (Castonguay, 1980; Lachapelle, 1980).

Deux stratégies d'analyse peuvent être utilisées pour étudier la mobilité linguistique: ou bien on s'intéresse aux conséquences qu'elle provoque sur des structures ou des phénomènes plus généraux, ou bien on s'efforce de mettre en évidence et de mesurer les processus qui l'engendrent ou les facteurs qui rendent compte de ses variations. Dans le premier cas, la mobilité linguistique représente une variable explicative ou, pour prendre le langage de la statistique, une variable indépendante. On s'est surtout intéressé, dans ce cadre d'analyse, aux effets qu'elle provoque sur la composition linguistique (Castonguay, 1983b; Castonguay et Marion, 1974; Lachapelle et Henripin, 1980). La mobilité linguistique est plutôt considérée, dans le second cas, comme une variable dépendante ou à expliquer. Certains ont adopté une approche classique (De Vries et Vallee, 1980; Gryz, 1980), que l'on qualifie parfois de factorielle. Elle consiste à présenter les variations du phénomène considéré en fonction de différentes variables (sexe, âge, état matrimonial, lieu de naissance, scolarité, lieu de résidence, etc.), prises une à une ou conjointement. Cette approche a l'avantage de ne reposer sur aucune d'hypothèse. Du fait de son caractère exploratoire, elle est bien adaptée à l'étude des phénomènes peu connus. D'autres ont préféré sauter cette étape et recourir d'entrée de jeu au langage des processus.

Ils ont montré qu'il est fructueux de concevoir la mobilité linguistique comme la résultante d'une série de sous-processus interdépendants et liés au cycle de vie: l'acquisition d'une langue seconde, le choix du conjoint et l'orientation linguistique chez les couples homoglottes et hétéroglottes (Castonguay, 1983a; Lachapelle, 1983). Si cette perspective permet souvent de conclure, elle repose toutefois sur de nombreuses hypothèses, plausibles certes mais difficiles à vérifier directement, en raison des limites des données censitaires.

Toutes ces stratégies et perspectives d'analyse ont en commun un même point de départ: une matrice ou un tableau de base dans lequel se trouve consignée, pour une sous-population donnée, la distribution conjointe de la langue maternelle et de la langue d'usage(2). On en tire des indices ad hoc qui mettent en lumière une partie de l'information contenue dans le tableau. Ceux-ci seront présentés dans la première section. D'autres indices découlent d'une analyse exhaustive ou sont définis par rapport à un modèle. La formalisation a notamment pour but de concilier les énoncés en apparence contradictoires qui découlent de l'examen de divers indices ad hoc. Elle confère aussi plus de cohérence à l'analyse. Cela conduit parfois à proposer des indices inédits; dans d'autres cas, le modèle donne une signification nouvelle à un indice fréquemment utilisé. Ces questions feront l'objet de la seconde section. Les mesures et les indices proposés seront illustrés à l'aide de données tirées des recensements du Canada. Les commentaires seront toujours fort brefs, car il ne s'agit pas ici d'examiner une situation réelle mais de fabriquer et de calibrer des outils d'analyse.

## Indices ad hoc

Désignons par $n_{ij}$ l'effectif des individus qui ont i pour langue d'origine (langue maternelle) et j pour langue d'arrivée (langue au foyer). On notera $n_{i.}$ l'effectif de la population ayant i pour langue maternelle et $n_{.i}$ le nombre des personnes qui ont i pour langue d'usage. On peut former de nombreux indicateurs à l'aide de ces informations. Certains font appel uniquement aux marges du tableau (les $n_{i.}$ et les $n_{.i}$), d'autre recourent aux éléments qui en constituent le noyau. Dans l'un et l'autre cas, il est possible de considérer soit des différences d'effectifs soit des rapports, des taux ou des propensions. Inventorions d'abord les indicateurs qui correspondent à des soldes nets ou qui en découlent.

### Différences d'effectifs

On peut classer les groupes linguistiques en deux catégories selon qu'il sont ou non avantagés par le processus de mobilité linguistique. Cela conduit à mesurer les transferts nets propres au groupe i, soit:

$$TN_i = n_{.i} - n_{i.} \qquad (I)$$

La somme de ces grandeurs est évidemment nulle, car:

$$n.. = \sum_i n_{.i} = \sum_i n_{i.} \qquad (II)$$

Pour que le processus ne profite à aucun groupe, il faut que les transferts nets soient tous nuls ou, en d'autres mots, que la composition par langue d'usage soit identique à la composition par langue maternelle.

Considérons les langues deux à deux. Les transferts nets que reçoit la langue i en provenance de la langue j correspondent à $TN_{ij}$ ($= n_{ji} - n_{ij}$). On obtient de cette manière une décomposition des transferts nets relatifs au groupe i, puisque:

$$TN_i = \sum_{j \neq i} TN_{ij} \qquad \text{(III)}$$

Les $TN_{ij}$ procurent aussi une indication sur l'ordre de dominance des langues en présence. Quand l'on observe tout à la fois $TN_{ij} > 0$, $TN_{jk} > 0$ et $TN_{ik} > 0$, il en résulte non seulement que la langue i domine la langue j et que celle-ci l'emporte sur la langue k, mais encore que la relation de dominance est transitive (la langue i l'emporte sur la langue k). Et si au surplus $TN_{ik} > TN_{jk}$, on pourra parler de transitivité complète.

Illustrons ces considérations à l'aide de données tirées du recensement du Canada effectué en 1981 (tableau 1). À en juger d'après les transferts linguistiques nets, l'anglais l'emporte partout sur le français et les deux langues officielles dominent les tierces langues. L'anglais est une langue dominante et le français une langue sous-dominante; quant aux tierces langues, elles sont dominées. Et comme prévu, la relation de dominance est complètement transitive. Ces énoncés s'appliquent aussi bien au Québec qu'au reste du pays, ce qui ne manque pas de surprendre, du moins à première vue.

Poussons plus loin l'analyse. Entre une langue dominante (i) et une langue dominée (k), les transferts linguistiques présentent une forte asymétrie, $n_{ik}$ étant négligeable par rapport à $n_{ki}$. Ce phénomène est mis en évidence en mesurant l'efficacité des échanges linguistiques, soit:

$$EF_{ik} = \max (n_{ik}, n_{ki}) \; / \; (n_{ik} + n_{ki}) \qquad \text{(IV)}$$

Cet indice prend ses valeurs entre 0.5 et 1. Quand les transferts linguistiques sont parfaitement équilibrés entre deux groupes, l'indice d'efficacité est égal à 0.5. Il atteint l'unité lorsque le groupe dominant n'effectue aucun transfert vers le groupe dominé. En première approximation, on peut considérer les transferts comme faiblement efficaces quand l'indice est inférieur à 0.75; ils seront tenus pour fortement efficaces s'ils dépassent ce seuil.

**TABLEAU 1**
**Transferts linguistiques nets entre les langues en présence, Canada, Québec, Canada moins le Québec, 1981[a]**

(En milliers[b])

| Région et langue | Transferts nets vers | | |
|---|---|---|---|
| | l'anglais | le français | les tierces langues |
| Canada | | | |
| Anglais | - | -288 | -1,387 |
| Français | 288 | - | -35 |
| Autre | 1,387 | 35 | - |
| Total | 1,675 | -253 | -1,422 |
| Québec | | | |
| Anglais | - | -24 | -90 |
| Français | 24 | - | -33 |
| Autre | 90 | 33 | - |
| Total | 114 | 8 | -123 |
| Canada moins Québec | | | |
| Anglais | - | -264 | -1,297 |
| Français | 264 | - | -3 |
| Autre | 1,297 | 3 | - |
| Total | 1,561 | -262 | -1,300 |

a Non compris les pensionnaires d'institution.
b En raison des arrondis, les totaux ne correspondent pas toujours à la somme de leurs éléments.

**Source:** Statistique Canada, Recensements du Canada de 1981; les calculs sont de l'auteur.

Au Québec, on l'a vu, l'anglais l'emporte sur le français et ces deux langues dominent nettement les tierces langues. Les indices d'efficacité confirment ce diagnostic (tableau 2). Les transferts entre l'anglais et le français se révèlent les moins efficaces (0.6); à l'opposé, les transferts entre l'anglais et les tierces langues présentent un haut degré d'efficacité (0.9). Comme prévu, l'efficacité des transferts entre le français et les tierces langues se situe à un niveau intermédiaire (0.8). Dans le reste du Canada, il y a cependant peu de différence entre l'efficacité des transferts d'une part entre l'anglais

et le français (près de 0.9) et d'autre part entre l'anglais et les tierces langues (un peu plus de 0.9). Cela donne à penser que le français y est une langue semi-dominée. Et c'est du reste ce que confirme la faible efficacité des transferts entre le français et les tierces langues (0.7).

**TABLEAU 2**
**Indice d'efficacité de transferts linguistiques, Canada, Québec et Canada moins le Québec, 1981[a]**

(En %)

| Région | Transferts entre | | |
|---|---|---|---|
| | l'anglais et le français | l'anglais et les tierces langues | le français et les tierces langues |
| Canada | 77 | 93 | 76 |
| Québec | 56 | 90 | 77 |
| Canada moins Québec | 88 | 93 | 66 |

[a] Non compris les pensionnaires d'institution.

**Source:** Statistique Canada, Recensement du Canada de 1981; les calculs sont de l'auteur.

D'autres indices corroborent ce diagnostic. C'est notamment le cas de la proportion des pertes nettes du tiers groupe qui vont au groupe anglais (le complément à un de cette proportion correspond à la part du groupe français). Elle s'élève à 73 % au Québec et à plus de 99 % ailleurs au Canada. On obtient des chiffres analogues quand on considère plutôt la fraction des transferts bruts en provenance des tierces langues qui se portent vers l'anglais (69 % au Québec et plus de 99 % dans le reste du Canada).

L'examen de quelques indices simples permet donc de ranger les langues en présence dans les deux grandes régions linguistiques du Canada

en quatre catégories: langue dominante (l'anglais partout), langue sous-dominante (le français au Québec), langue semi-dominée (le français ailleurs au Canada) et langue dominée (les tierces langues partout). Ce classement correspond-il à une structure fondamentale ou découle-il seulement des indices choisis?

**Rapports d'effectifs**

Au lieu d'examiner les effectifs observés, il est souvent préférable, pour faciliter les comparaisons, de considérer des effectifs réduits ou des proportions. Notons $p_{ij}$ la fraction de la population qui a i pour langue maternelle et j pour langue au foyer. Il en résulte immédiatement que:

$$p_{i.} = \sum_j p_{ij}$$

$$p_{.i} = \sum_j p_{ji}$$

$$\sum_i p_{i.} = \sum_i p_{.i} = 1 \qquad \text{(V)}$$

où $p_{i.}$ et $p_{.i}$ désignent respectivement la proportion de la population qui a i soit pour langue maternelle soit pour langue au foyer. On peut évidemment reformuler les indices que nous venons de présenter à partir des effectifs réduits. Cela n'entraîne aucun changement dans le signe ni dans le sens des différences. Même l'indice d'efficacité n'est pas modifié par cette transformation.

S'agissant d'apprécier l'effet net des transferts sur un groupe linguistique, on a souvent utilisé le taux de transfert net (Rochon-Lesage et Maheu, 1974):

$$tn_i = (p_{.i} - p_{i.})/p_{i.} = (n_{.i} - n_{i.})/n_{i.}$$

$$= TN_i/n_{i.} \qquad \text{(VI)}$$

Ce taux est positif quand un groupe profite des transferts linguistiques et négatif dans le cas contraire. Si l'on s'intéresse seulement aux groupes linguistiques qui font des pertes, il semble naturel de considérer plutôt l'opposé de ce taux (Castonguay et Marion, 1974), soit le taux net de mobilité linguistique ($tm_i$). Une autre mesure a aussi été employée, l'indice de continuité linguistique (Lachapelle et Henripin, 1980). Il se définit comme suit:

$$c_i = n_{.i}/n_{i.} = p_{.i}/p_{i.} \qquad \text{(VII)}$$

Ces trois mesures ont entre elles des relations fort simples:

$$tm_i = - tn_i$$
$$\text{et} \quad tm_i = 1 - c_i \qquad \text{(VIII)}$$

Elles devraient donc conduire toujours aux mêmes conclusions. C'est dire que, selon les besoins de l'analyse, on peut choisir l'une ou l'autre de ces grandeurs pour mesurer l'effet net de la mobilité linguistique. Encore faut-il tenir compte des risques d'erreur que recèle chacune d'elles.

Tout indice représente, on le sait, une valeur approchée de la grandeur réelle qu'il entend mesurer. Notons $e_i$ l'erreur de mesure qui affecte $tn_i$. Cela permet d'écrire, en vertu de l'équation précédente:

$$tn_i + e_i = c_i + e_i - 1 \qquad \text{(IX)}$$

L'erreur faite dans l'estimation du taux de transfert net est d'importance égale à celle qui affecte l'indice de continuité. Mais il en va autrement quand on considère la valeur absolue de l'erreur relative (er). En effet:

$$er\ (tn_i) = \left| e_i/tn_i \right| = er\ (tm_i)$$

et

$$er\ (c_i) = \left| e_i/c_i \right| \qquad \text{(X)}$$

Il en découle immédiatement que:

$$er\ (c_i)\ /\ er(^{tm}_i) = |\ ^{tm}_i\ |\ /\ c_i \qquad (XI)$$

L'erreur relative qui affecte les taux nets sera égale à celle qui s'applique à l'indice de continuité seulement quand $c_i = {}^{tm}_i = -{}^{tn}_i = 0.5$. Lorsque l'indice de continuité est supérieur à 0.5, l'erreur relative du taux net pourra être nettement plus élevée que celle qui affecte l'indice de continuité. La situation est évidemment renversée quand l'indice de continuité est inférieur à 0.5.

Donnons un exemple pour y voir plus clair. Dans une population donnée, l'indice de continuité observé chez un groupe linguistique correspond à 1.001. On commet ce faisant une erreur de -0.002, l'indice réel valant 0.999. L'erreur relative s'élève donc à 2 pour 1000 (0.002/1.001). Elle est toutefois beaucoup plus élevée pour le taux de transfert linguistique net. Certes, la différence entre sa valeur observée (0.001) et sa valeur réelle (-0.001) est identique à celle qui affecte l'indice de continuité, mais l'erreur relative (0.002/0.001) est mille fois plus forte.

Faut-il dès lors adopter dans chaque cas la mesure qui est affectée de la plus faible erreur relative? Une pareille consigne serait sans doute excessive. D'ailleurs elle compliquerait inutilement l'analyse, la mesure retenue pouvant varier d'un groupe linguistique ou d'une population à l'autre. Il n'en demeure pas moins nécessaire, dans l'interprétation des écarts, de toujours avoir présent à l'esprit les rapports entre les erreurs relatives qui affectent chacune des mesures.

Quand un groupe est avantagé par le processus de mobilité linguistique, son taux de transfert net est positif et la valeur de l'indice de continuité dépasse l'unité. Partout au Canada l'indice de continuité du groupe anglais est supérieur à celui du groupe français (tableau 3); quant à l'indice de continuité du tiers groupe, il est toujours inférieur à celui du groupe français. À noter toutefois qu'au Québec l'indice du groupe français excède légèrement l'unité tandis qu'ailleurs au Canada il

**TABLEAU 3**
**Indices ad hoc permettant de décrire la situation linguistique, Canada, Québec et Canada moins Québec, 1981[a]**

(En %)

| | Composition de la population selon la | | Taux de transfert net | Taux net de mobilité linguistique | Indice de continuité | Valeur réduite des transferts linguistiques nets[d] |
|---|---|---|---|---|---|---|
| | langue maternelle[b] | langue parlée à la maison | | | | |
| Canada | | | | | | |
| Anglais | 61.3 | 68.2 | 11.4 | -11.4 | 111.4 | 7.0 |
| Français | 25.7 | 24.6 | -4.1 | 4.1 | 95.9 | -1.1 |
| Autre | 13.0 | 7.2 | -45.1 | 45.1 | 54.9 | -5.9 |
| (Effectif total[c]) | (24,343) | (24,083) | | | | |
| Québec | | | | | | |
| Anglais | 11.0 | 12.7 | 16.4 | -16.4 | 116.4 | 1.8 |
| Français | 82.4 | 82.5 | 0.2 | -0.2 | 100.2 | 0.1 |
| Autre | 6.6 | 4.8 | -28.8 | 28.8 | 71.2 | -1.9 |
| (Effectif total[c]) | (6,438) | (6,369) | | | | |
| Canada moins Québec | 79.4 | 88.2 | 11.1 | -11.1 | 111.1 | 8.8 |
| Anglais | 5.3 | 3.8 | -28.2 | 28.2 | 71.8 | -1.5 |
| Français | 15.4 | 8.1 | -47.6 | 47.6 | 52.4 | -7.3 |
| Autre | | | | | | |
| (Effectif total[c]) | (17,905) | (17,714) | | | | |

a Non compris les pensionnaires d'institution.
b Y compris, dans cette seule colonne, les pensionnaires d'institution.
c En milliers.
d Cet indice correspond à $p_{\cdot i} - p_{i\cdot}$ dans la population considérée (non compris les pensionnaires d'institution).

**Source:** Statistique Canada, Recensement du Canada de 1981; les calculs sont de l'auteur.

est nettement inférieur à cette valeur (0.7). Ces constations s'accordent avec celles qui ressortent de l'examen des différences entre les effectifs des transferts linguistiques.

L'indice de continuité représente-il par ailleurs une mesure de l'attraction relative qu'exercent les différentes langues en présence? Une pareille interprétation ne va pas sans difficulté, car on voit mal ce qui permettrait d'avancer que l'anglais possède une force d'attraction plus grande au Québec (1.16) qu'ailleurs au Canada (1.11) Notons à ce propos que la plage de variation de l'indice de continuité dépend de la composition linguistique de départ (ou par langue maternelle). Il est toujours compris, pour un groupe linguistique donné, entre 0 et $1/p_i$. Pour le groupe anglais, l'indice de continuité a donc au Québec une valeur maximale de l'ordre de 9, tandis qu'ailleurs au Canada il ne pourrait excéder 1.26, ce qui rend délicate l'interprétation des écarts observés. Les valeurs réduites des transferts linguistiques nets (tableau 3) montrent bien du reste que la mobilité linguistique procure des gains plus importants au groupe anglais à l'extérieur du Québec que dans cette province.

On a proposé d'utiliser les taux de mobilité linguistique pour mesurer l'intensité du phénomène, pour une langue donnée, dans différentes populations (Lachapelle et Henripin, 1980). Notons $t_{ij}$ le taux de mobilité linguistique de la langue de départ i à la langue d'arrivée j. La valeur prise par ce taux correspond à:

$$t_{ij} = n_{ij}/n_{i.} = p_{ij}/p_{i.} \qquad \text{(XII)}$$

Quand i = j, on parlera plutôt de taux de persistance linguistique. Quelle que soit la langue i considérée, on a toujours:

$$\sum_j t_{ij} = 1 \qquad \text{(XIII)}$$

Le taux de persistance et son complément à 1 (le taux global de mobilité linguistique) ne sont pas sans rapport avec les grandeurs qui ont été présentées plus haut. Il est en effet facile de démontrer que:

$$c_i \geq p_{ii}$$

et que

$$tm_i \leq 1-p_{ii} \qquad \text{(XIV)}$$

L'indice de continuité est une approximation par excès du taux de persistance tandis que le taux net de mobilité linguistique procure une estimation par défaut du taux global de mobilité linguistique. L'approximation sera souvent satisfaisante pour les langues dominées qui reçoivent peu de transferts en provenance des autres langues. Cette remarque vaut surtout au Canada pour le tiers groupe ainsi que pour le groupe français à l'extérieur du Québec.

On a également suggéré d'utiliser les taux de mobilité linguistique pour caractériser la situation respective des langues en présence au sein d'une population (Kralt, 1977). Si la prise en considération de ces nouveaux indices ne modifie en rien le diagnostic qui a été fait pour l'ensemble des territoires et provinces autres que le Québec, on serait cependant amené à conclure, dans le cas de cette province, - et contre toute attente - à une domination du français sur l'anglais (tableau 4), le taux de mobilité de l'anglais au français (12 %) étant nettement supérieur à celui qui est observé du français à l'anglais (2 %). Ce paradoxe n'a rien de nouveau: il a également été relevé dans les résultats du recensement de 1971 (Castonguay, 1980; Lachapelle, 1980). Il mérite d'être examiné de plus près.

**TABLEAU 4**
**Taux de persistance et de mobilité linguistiques, Canada, Québec et Canada moins le Québec, 1981[a]**

(En %)

| Région et langue maternelle | | Langue parlée à la maison | | |
|---|---|---|---|---|
| | | Anglais | Français | Autre |
| Canada | | | | |
| | Anglais | 98.4 | 0.8 | 0.8 |
| | Français | 6.6 | 93.1 | 0.3 |
| | Autre | 47.4 | 1.6 | 51.0 |
| Québec | | | | |
| | Anglais | 86.5 | 11.8 | 1.7 |
| | Français | 2.0 | 97.7 | 0.3 |
| | Autre | 23.9 | 10.9 | 65.2 |
| Canada moins Québec | | | | |
| | Anglais | 99.0 | 0.3 | 0.7 |
| | Français | 32.8 | 66.9 | 0.3 |
| | Autre | 51.1 | 0.2 | 48.7 |

[a] Non compris les pensionnaires d'institution.

**Source:** Statistique Canada, Recensement du Canada de 1981; les calculs sont de l'auteur.

## Indices dérivés de modèles

Le nombre d'indices ad hoc n'est borné en définitive que par l'imagination des chercheurs. Cette surabondance d'outils d'analyse est cependant illusoire, car il n'y a pas un nombre illimité de paramètres indépendants. Un tableau croisé comportant n langues maternelles et n langues parlées au foyer renferme en effet $n^2$ degrés de liberté ($n^2 - 1$ quand l'on s'en tient aux effectifs réduits). C'est dire que l'on ne peut définir plus de $n^2$ paramètres indépendants. C'est là une contrainte générale. Elle n'implique évidemment pas que l'on puisse déterminer les meilleurs indices possibles. Cela dépend du point de vue retenu.

S'agissant d'étudier la mobilité linguistique, deux perspectives viennent immédiatement à l'esprit. Ou bien on se préoccupe surtout des effets structurels du phénomène et l'on porte une attention particulière aux marges du tableau (les compositions selon la langue maternelle et selon la langue parlée au foyer), ou bien l'on s'intéresse plutôt aux propensions et on examine d'abord et avant tout les taux de mobilité. On développera tour à tour chacune de ces perspectives de manière détaillée. Pour simplifier l'exposé, on supposera que seulement deux langues sont en présence (i et j).

### De la composition linguistique aux taux de mobilité

Supposons que l'on connaisse la composition par langue maternelle (les $n_{i.}$ et les $P_{i.}$) ainsi que la composition par langue d'usage (les $n_{.i}$ et les $P_{.i}$), mais que l'on ignore l'effectif des transferts de i à j ($n_{ij}$ et $P_{ij}$) et de j à i ($n_{ji}$ et $p_{ji}$). S'il n'est pas possible dans ces conditions de préciser la valeur des taux de mobilité linguistique ($t_{ij}$ et $t_{ji}$), il n'en demeure pas moins que la plage de variation de ces taux n'est pas indépendante des compositions linguistiques observées.

Admettons, sans perte de généralité, que le groupe i est toujours gagnant dans les échanges linguistiques ou en tout cas qu'il n'y perd

jamais. Cela revient à supposer que $c_i \geq 1 \geq c_j$, ce qui est équivalent à $P_{.i} \geq P_i$. Sachant que $P_{ji} - P_{ij} = P_{.i} - P_{i.}$, ces hypothèses impliquent que:

$$P_{ji} \geq P_{ij}$$

et que

$$(1-P_{i.})\ t_{ji} \geq P_{i.}\ t_{ij} \qquad \text{(XV)}$$

Cette dernière inégalité indique clairement que les taux de mobilité ne varient pas indépendamment de la composition linguistique de départ.

Pour y voir plus clair, considérons en premier lieu le cas de figure selon lequel le groupe gagnant (i) est également majoritaire dans la population ($P_{i.} \geq 1-P_{i.}$). Il vient (d'après XV):

$$t_{ji}/t_{ij} \geq P_{i.}/(1 - P_{i.}) \geq 1 \qquad \text{(XVI)}$$

Le taux de mobilité du groupe minoritaire est toujours supérieur à celui du groupe majoritaire. Il y a toutefois une exception, mais elle est triviale. Quand $c_i = c_j = 1$, il est possible que les transferts linguistiques de même que les taux de mobilité soient tous nuls. Et si au surplus les groupes linguistiques sont d'égale importance, il va de soi que les taux de mobilité, quelle qu'en soit l'intensité, sont égaux.

En raison des contraintes qu'imposent les compositions linguistiques, les taux de mobilité ne varient pas entre 0 et 1. Il est facile de montrer que le taux du groupe majoritaire ($t_{ij}$) se situera toujours entre 0 et $P_{.j}/P_i$. tandis que celui du groupe minoritaire ($t_{ji}$) prendra ses valeurs entre $^{tm}_{j}$ et 1, où $^{tm}_{j}$ désigne, rappelons-le, le taux net de mobilité du groupe j.

Il paraît naturel de normaliser les taux de mobilité pour qu'ils aient une même plage de variation et que celle-ci soit comprise entre 0 et 1. Si l'on convient d'utiliser le symbole g ($g_{ij}$ et $g_{ji}$) pour désigner les taux normalisés, on a:

$$g_{ij} = t_{ij} \; / \; p_{.j}/p_{i.} = p_{ij}/p_{.j}$$

et $$g_{ji} = (t_{ji}-tm_j) \; / \; (1-tm_j)$$

$$= (p_{ji} - (p_{.i}-p_{i.})) \; / \; 1-p_{.i}$$

$$= p_{ij}/p_{.j} = g_{ij} \qquad \text{(XVII)}$$

Les taux normalisés sont toujours égaux entre eux. Le paramètre g (= $p_{ij}/p_{.j}$) caractérise la mobilité linguistique qui intervient entre les deux groupes en présence. On peut d'ailleurs exprimer la valeur des taux de mobilité en fonction de ce paramètre:

$$t_{ij} = g \; (1 - p_{.i}) \; / \; p_{i.} \; = \; g(1-tm_j) \; (1-p_{i.})/p_{i.}$$

et $$t_{ji} = tm_j + g \; (1-tm_j)$$

ou encore $$t_{ji} = tm_j + t_{ij} \; p_{i.} \; / \; (1 - p_{i.}) \qquad \text{(XVIII)}$$

Quand g est nul, le groupe majoritaire et dominant a un taux de mobilité nul; et celui du groupe minoritaire s'abaisse à sa valeur minimale ($tm_j$). À l'inverse, quand g vaut 1, les taux atteignent tous deux leur valeur maximale.

Outre les termes qui correspondent aux éléments des compositions linguistiques, deux paramètres apparaîssent dans les équations précédentes, g et $tm_j$. On sait que $tm_j$ mesure le déséquilibre de transferts au détriment du groupe j. Peut-on interpréter g comme une mesure de l'intensité des transferts équilibrés? Si $tm_j$ est nul, on a, bien sûr, $p_{i.} = p_{.i}$. Il vient:

$$t_{ij} = g \; (1-p_{i.})/p_{i.}$$

$$t_{ji} = g \qquad \text{(XIX)}$$

Et cela implique que les transferts linguistiques sont équilibrés, car:

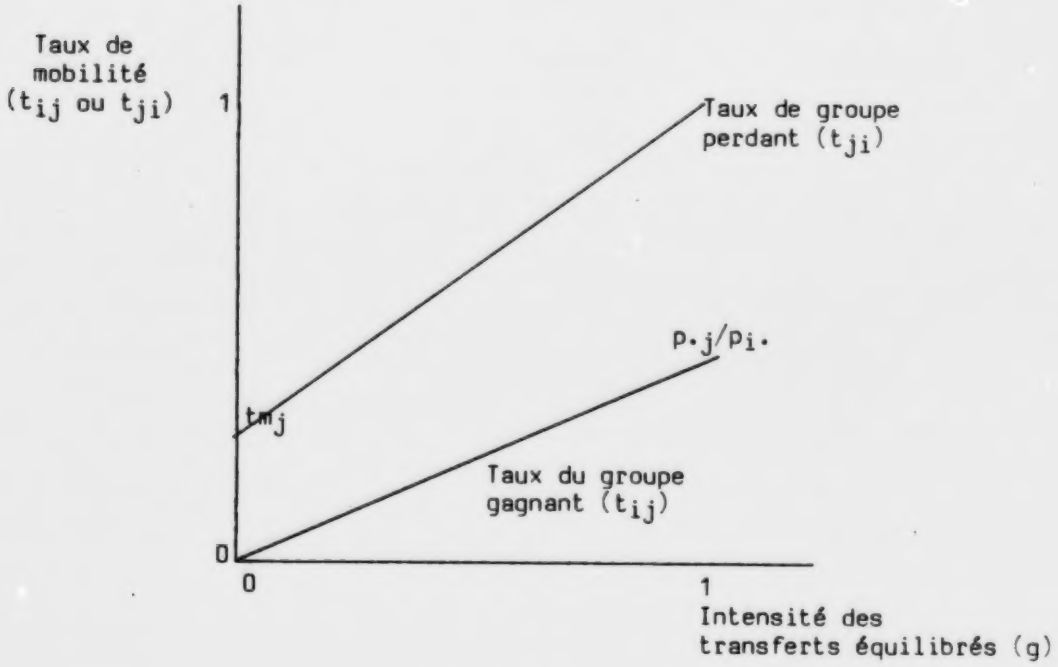
$$p_{ij} = p_{i.} \; t_{ij} = g \; (1-p_{i.})$$

et $$p_{ji} = (1-p_{i.}) \; t_{ji} = g \; (1-p_{i.}) \qquad \text{(XX)}$$

Trois paramètres indépendants épuisent la totalité de l'information contenue dans le tableau croisé (langue maternelle par langue au foyer): 1° le poids du groupe gagnant dans la population ($P_{i.}$) qui décrit la composition par langue maternelle; 2° le taux net de mobilité du groupe perdant ($tm_j$) qui mesure le déséquilibre des transferts linguistiques; et 3° un coefficient (g) qui correspond à l'intensité des échanges équilibrés entre les groupes en présence.

Il y a (en vertu de XVIII) une relation linéaire entre les taux de mobilité et le coefficient g (figure 1). Ces droites sont parfaitement

**Figure 1.** Relation entre les taux de mobilité et les paramètres fondamentaux quand le groupe gagnant est majoritaire dans la population




définies quand on connaît les compositions linguistiques ou, ce qui revient du même, les paramètres $P_{i.}$ et $tm_j$. La valeur que prend g permet de préciser l'intensité des taux de mobilité dans la population considérée.

Peut-on établir des relations analogues dans le cas où le groupe dominant est minoritaire ($P_{i.} \leq 1-P_{i.}$)? Essayons d'abord de déterminer les bornes de l'intervalle de variation des taux de mobilité. Le groupe i étant minoritaire, son taux de mobilité ($t_{ij}$) peut prendre toutes les valeurs comprises entre 0 et 1. Le taux de mobilité du groupe majoritaire ($t_{ji}$) possède toutefois une plage de variation plus étroite: il a en effet $tm_j$ pour borne inférieure et $P_{.i}/(1-P_{i.})$ comme borne supérieure. Pour établir ces résultats, il a fallu admettre que $P_{j.} > P_{.i}$. Cette condition n'est pas toujours satisfaite, en particulier lorsque le groupe dominant représente près de la moitié de la population(3).

Quelle est la valeur du coefficient g? Par un raisonnement analogue à celui auquel nous avons fait appel plus haut, on établit facilement qu'il correspond à:

$$g_{ij} = t_{ij}$$

et
$$g_{ji} = (t_{ji}-tm_j) \,/\, ((P_{.i}/1-P_{i.}) - tm_j)$$
$$= (P_{ji} - (P_{.i}-P_{i.})) \,/\, P_{i.}$$
$$= P_{ij}/P_{i.} = g_{ij} \qquad \text{(XXI)}$$

Bien que g ait le même numérateur que dans l'équation (XVII), le dénominateur est cependant différent. En fait, quelle que soit la situation considérée, g correspond au rapport du plus faible nombre de transferts linguistiques au plus petit effectif porté par l'une ou l'autre des marges du tableau, soit:

$$g = \min (P_{ij},P_{ji}) \,/\, \min (P_{i.},P_{j.},P_{.i},P_{.j}) \qquad \text{(XXII)}$$

Le complément à un de g correspond évidemment a un indice général de persistance ou d'immobilité (Boudon, 1973).

Il reste maintenant à exprimer, à partir de l'équation (XXI), la valeur des taux de mobilité en fonction de g:

$$t_{ij} = g$$

et $$t_{ji} = {}^{t}m_j + g\, p_{i.}/(1-p_{i.})$$

ou bien: $$t_{ji} = {}^{t}m_j + t_{ij}\, p_{i.}/(1-p_{i.}) \qquad \text{(XXIII)}$$

La dernière égalité est identique à celle qui apparaît en XVIII. Dans ce dernier cas, on avait toujours $t_{ji} > t_{ij}$ parce que le groupe dominant était majoritaire. Qu'en est-il lorsque ce groupe est minoritaire? Établissons d'abord la valeur de a tel que $t_{ij} = t_{ji} = a$. Il vient, en vertu de la troisième égalité en XXIII:

$$a - {}^{t}m_j = a\, p_{i.}/(1-p_{i.})$$

$$a\,(1-2p_{i.})/(1-p_{i.}) = {}^{t}m_j$$

$$a = (p_{.i}-p_{i.})/(1-2p_{i.}) \qquad \text{(XXIV)}$$

Ce coefficient est inférieur ou égal à un, à condition que $p_{j.} \geq p_{.i}$. Cette condition peut ne pas être satisfaite lorsque le groupe dominant représente une très forte minorité. Cette situation se ramène, on peut le démontrer, au cas où le groupe dominant est majoritaire. Notons par ailleurs que le coefficient a est toujours supérieur à ${}^{t}m_j$.

A quelle condition peut-on observer $t_{ij} > t_{ji}$? Cela équivaut à écrire:

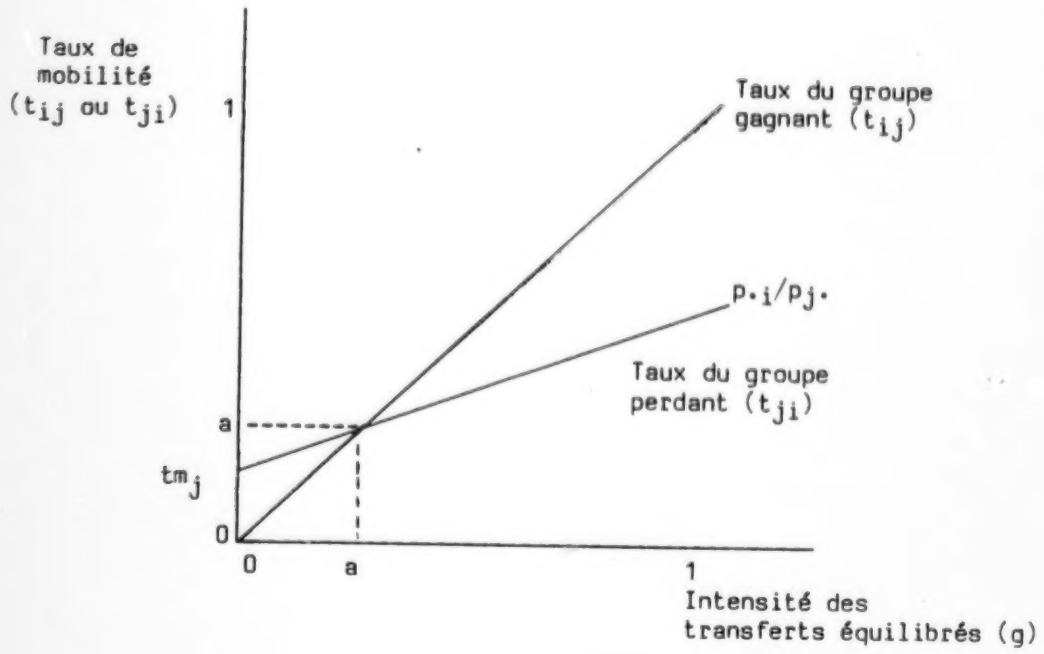
$$t_{ij} > {}^{t}m_j + t_{ij}\, p_{i.} \,/\, (1-p_{i.})$$

$$t_{ij}\,(1-2p_{i.}) > {}^{t}m_j\,(1-p_{i.})$$

$$t_{ij} > (p_{.i}-p_{i.}) \,/\, (1-2p_{i.}) = a \qquad \text{(XXV)}$$

On vérifie aussi que $t_{ji} > a$. En d'autres mots, quand g a, il vient $t_{ij} > t_{ji} > a$; symétriquement, lorsque g <a, on a toujours $a > t_{ji} > t_{ij}$. Les droites qui portent les taux de mobilité se croisent quand g prend la valeur a (figure 2).

**Figure 2.** Relation entre les taux de mobilité et les paramètres fondamentaux quand le groupe gagnant est minoritaire dans la population




La situation observée au Québec entre les groupes anglais et français n'a donc rien de paradoxal. Quand le groupe gagnant est minoritaire, elle survient chaque fois que le paramètre g a une valeur supérieure à celle du coefficient a.

Résumons-nous. On connaît souvent les marges d'un tableau tout en ignorant son contenu exact. On constate à l'examen de la composition par langue maternelle ($P_{i\cdot}$ et $P_{j\cdot}$) et de la composition par langue d'usage ($P_{\cdot i}$ et $P_{\cdot j}$) que le groupe i l'emporte sur le groupe j ($P_{ji} - P_{ij} = P_{\cdot i} - P_{i\cdot} > 0$). Compte tenu de cette information, que peut-on dire sur les taux de mobilité de chaque groupe ($t_{ij}$ et $t_{ji}$)? On a établi que le taux de mobilité du groupe perdant ($t_{ji}$) est toujours supérieur à celui du groupe gagnant ($t_{ij}$) quand ce dernier est majoritaire dans la population. Cette situation se rencontre aussi lorsque le groupe gagnant est

minoritaire, à condition toutefois que les deux taux de mobilité soient inférieurs à un certain seuil (a), lequel dépend uniquement des marges du tableau. Au-delà de ce seuil, la situation se modifie et on constate que le taux de mobilité du groupe gagnant ($t_{ij}$) dépasse celui du groupe perdant ($t_{ji}$).

On a mis en évidence, chemin faisant, deux paramètres fondamentaux: le premier, en l'espèce le taux net de mobilité linguistique du groupe perdant ($tm_j$), mesure l'effet structurel que provoque la mobilité linguistique; il tire parti, avec la proportion du groupe gagnant dans la population de départ ($P_{i.}$), de toute l'information que renferment les marges du tableau. Le second coefficient (g) mesure l'intensité des transferts linguistiques équilibrés. Il épuise l'information contenue dans le tableau: un tableau croisé à deux catégories dont on connaît les marges n'a en effet qu'un degré de liberté.

Donnons un exemple. D'après le recensement de 1981, il y a au Québec 601,555 personnes qui ont l'anglais pour langue maternelle et pour langue d'usage; 82,135 personnes font un transfert de l'anglais au français; les transferts du français à l'anglais sont au nombre de 106,370 alors que 5,128,135 personnes ont le français pour langue maternelle et pour langue d'usage. On voit que les transferts linguistiques favorisent nettement le groupe anglais. Le taux net de mobilité linguistique du groupe français, ou taux net d'anglicisation ($tm_j$), vaut d'ailleurs +0.46 %. On constate toutefois que le taux de mobilité du groupe anglais (12.0 %), qui se confond ici avec la mesure (g) de l'intensité des transferts équilibrés, dépasse de loin celui du groupe français (2.0 %). C'est que ces taux sont de loin supérieurs au seuil que détermine le coefficient a, dont la valeur est ici très faible (0.53 %). Convenons de désigner par la lettre i le groupe anglais et par la lettre j le groupe français. On peut développer comme suit la dernière égalité apparaissant en XXIII:

$$p_{ji} = (p_{.i} - p_{i.}) + p_{ij}$$

et $$n_{ji} = (n_{.i} - n_{i.}) + n_{ij} \qquad \text{(XXVI)}$$

Le premier terme du membre de droite correspond à l'effet structurel des transferts linguistiques; les transferts équilibrés sont saisis par le second terme. On a dans notre exemple:

$$106{,}370 = 24{,}235 + 82{,}135$$

Si l'on modifie la valeur de g, qui vaut ici 12.0 %, l'effet structurel ne sera évidemment pas modifié. Adoptons la valeur que prend le coefficient a (0.53 %). On obtient:

$$27{,}876 = 24{,}235 + 3{,}641$$

La part des transferts équilibrés est passée de 77 % (82,135/106,370) à 13 % (3,641/27,876). Il eut donc fallu une très forte baisse du nombre de transferts pour que les taux de mobilité soient égaux entre eux (à 0.53 %). Et elle aurait dû être encore plus forte pour renverser l'ordre des taux de mobilité. La situation observée au Québec en ce qui a trait à la mobilité linguistique entre les groupes anglais et français n'a donc rien d'étonnant. On observe du reste une situation analogue ailleurs au Canada entre le groupe français et le tiers groupe (tableaux 1 et 4).

Pour résoudre ce problème, Castonguay (1980) a proposé une approche différente de celle que nous venons de présenter. Elle consiste à supposer que les taux de mobilité sont proportionnels à l'importance de la langue de destination dans la population de départ, soit:

$$t_{ij} = b_{ij}\, p_{j.}$$

et $$t_{ji} = b_{ji}\, p_{i.} \qquad \text{(XXVII)}$$

Il paraît dès lors naturel d'assimiler les coefficients b à des taux normalisés de mobilité. Il valent respectivement:

$$b_{ij} = p_{ij} / p_i . p_j .$$

et

$$b_{ji} = p_{ji} / p_i . p_j . \qquad \text{(XXVIII)}$$

Ces taux ne prennent pas en considération les modifications structurelles qu'opère la mobilité linguistique. C'est dire qu'il peut y avoir des différences importantes dans les intervalles de variation de ces taux, ce qui ne facilite pas les comparaisons.

L'approche que nous avons développée peut se généraliser à un nombre quelconque (n) de langues. Il suffit en effet de mettre en évidence trois séries de paramètres indépendants:

1° la composition par langue maternelle (n-1 paramètres);
2° l'effet net des transferts entre chaque couple de langues (n(n-1)/2 coefficients);
3° l'intensité de la mobilité équilibrée entre chaque couple de langues (n(n-1)/2 paramètres).

Donnons quelques indications sur le mode de calcul des paramètres non triviaux. La première étape consiste à déterminer, pour chaque couple de langues, le groupe gagnant et le groupe perdant. Supposons, sans perte de généralité, que le groupe i l'emporte sur le groupe j. Posons $^{m}_{ji} = p_{ji} - p_{ij}$ et $^{m}_{ij} = 0$. Cela permet de calculer la valeur minimale de chaque taux de mobilité, que l'on notera $^{tm}_{ij}$ ($= m_{ij}/p_i .$). Pour chaque couple de langues, le plus élevé de ces taux nets de mobilité (il y en a n(n-1)/2) mesure l'effet net des transferts linguistiques. Il est un peu plus difficile d'établir la valeur des paramètres g (au nombre de n(n-1)/2). Sans entrer dans le détail des raisonnements, on vérifiera facilement que l'on a toujours:

$$g_{ij} = \min (p_{ij}, p_{ji}) / \min \left(p_{\cdot i} - \sum_{l \neq i} m_{li},\ p_{\cdot j} - \sum_{l \neq j} m_{lj}\right)$$

$$= \min (p_{ij}, p_{ji}) / \min \left(p_{i \cdot} - \sum_{l \neq i} m_{il},\ p_{j \cdot} - \sum_{l \neq j} m_{jl}\right) \qquad \text{(XXIX)}$$

Cela implique que $g_{ij} = g_{ji}$, comme on pouvait s'y attendre.

Ces indices permettent de bien distinguer deux types de mobilité linguistique: 1° une mobilité que l'on peut qualifier de structurelle parce qu'elle a une incidence sur la structure démolinguistique; et 2° une mobilité d'échange qui mesure l'intensité des transferts linguistiques équilibrés. Il s'agit là de deux dimensions fort importantes du phénomène; l'approche présentée y donne accès. Mais ce n'est évidemment pas la seule méthode d'analyse disponible.

**Des taux de mobilité à la composition linguistique**

Quand on connaît seulement les taux de mobilité ($t_{ij}$ et $t_{ji}$) et les taux de persistance ($t_{ii}$ et $t_{jj}$), que peut-on dire de la composition par langue maternelle et de la composition par langue d'usage? En l'absence d'information supplémentaire, supposons que la population est à l'état d'équilibre, ce qui revient à admettre que $P_{i.} = P_{.i}$ et que $P_{ji} = P_{ij}$. Désignons par $p^*_i$ et $p^*_j$ ($= 1-p^*_i$) cette composition d'équilibre. Il vient:

$$p^*_i \, t_{ij} = (1-p^*_i) \, t_{ji}$$

$$p^*_i = t_{ji}/(t_{ij} + t_{ji})$$

$$p^*_j = t_{ij}/(t_{ij} + t_{ji}) \qquad \text{(XXX)}$$

Ces relations ne permettent pas de déterminer lequel des deux groupes sort gagnant des échanges linguistiques réels. Tout ce qu'on peut avancer, c'est que si $t_{ji} \geq t_{ij}$, on a $p^*_i \geq p^*_j$ et que, symétriquement, si $t_{ij} > t_{ji}$, il vient $p^*_j > p^*_i$. En d'autres mots, le groupe i est majoritaire dans la composition d'équilibre lorsque son taux de mobilité est inférieur à celui de l'autre groupe.

Pour y voir plus clair, comparons la composition d'équilibre aux compositions observées. Afin de faciliter les développements, posons:

$$v = P_{ji} - P_{ij} = P_{.i} - P_{i.}$$

$$k = t_{ij} + t_{ji}$$

et $w = v/k$ (XXXI)

On a toujours:

$$v = (1-P_{i.})\, t_{ji} - P_{i.}\, t_{ij}$$

En remplaçant les taux de mobilité par leur valeur en XXX, on obtient:

$$v = (1-P_{i.})\, p_i^*\, k - P_{i.}\, (1-p_i^*)\, k$$

$$v = k\, (p_i^* - P_{i.})$$

et enfin $w = p_i^* - P_{i.}$ (XXXII)

Comme k est nul ou positif, il en résulte que v est toujours du même signe que w. Cela implique que si i est un groupe gagnant, son poids relatif est plus élevé dans la composition d'équilibre que dans la composition par langue maternelle. On observe évidemment l'inverse quand i désigne un groupe perdant.

Il reste à comparer la composition d'équilibre à la composition par langue d'usage. Posons $u = p_i^* - P_{.i}$ et remarquons que:

$$p_i^* - P_{.i} = (p_i^* - P_{i.}) - (P_{.i} - P_{i.})$$

$$u = w - v$$

$$= w\,(1-k) \qquad \text{(XXXIII)}$$

Si l'on admet que le coefficient k (= $t_{ij} + t_{ji}$) est inférieur à un, il en résulte que u est toujours du même signe que w. Sauf circonstances tout à fait spéciales ($k > 1$) qui ne risquent pas d'être rencontrées dans l'étude de la mobilité linguistique, il n'y a finalement que trois possibilités à considérer: a) le groupe i est gagnant ($p^*_i > p_{\cdot i} > p_{i\cdot}$), b) les deux groupes en présence sont en situation d'équilibre ($p^*_i = p_{i\cdot} = p_{\cdot i}$) et c) le groupe i est perdant ($p_{i\cdot} > p_{\cdot i} > p^*_i$).

Quand le groupe gagnant (i) est en situation minoritaire, il est évidemment possible que son taux de mobilité ($t_{ij}$) soit plus élevé que celui du groupe perdant ($t_{ji}$). Les résultats précédents permettent toutefois d'affirmer que l'on aura toujours:

$$t_{ji}/(t_{ij} + t_{ji}) = p^*_i > p_{\cdot i} > p_{i\cdot} \qquad \text{(XXXIV)}$$

En définitive, la composition linguistique d'équilibre n'est rien d'autre qu'un réaménagement des taux de mobilité.

Il est relativement aisé de généraliser cette méthode à un nombre quelconque de langues. Il suffit en effet de recourir à la théorie des chaînes de Markov finies (Bernard, 1978; Boudon, 1973; Gordon, 1965; Kemeny, Snell et Thompson, 1965). On peut présenter ce modèle de deux manières différentes. Supposons que l'on dispose du tableau des échanges linguistiques entre n langues. À partir de ces données, on calcule n(n-1) taux de mobilité et n taux de persistance (ceux-ci se déduisent de ceux-là). La théorie des chaînes de Markov stationnaires permet d'affirmer qu'il existe un état d'équilibre stable tel que la composition par langue maternelle est identique à la composition par langue d'usage. La composition d'équilibre dépend uniquement des taux de mobilité. En la comparant aux compositions observées, on peut le plus souvent distinguer les groupes gagnants des groupes perdants. C'est la méthode que nous avons suivie jusqu'ici.

S'agissant d'étudier la mobilité linguistique au Canada, on considère habituellement trois catégories de langue: l'anglais, le français et une catégorie résiduelle, les tierces langues. Examinons, à titre illustratif, l'écart entre la composition d'équilibre et les compositions linguistiques observées en 1971 dans la population du Québec appartenant au groupe d'âges 25-44 ans:

| | Par langue maternelle | Par langue d'usage | À l'état d'équilibre |
|---|---|---|---|
| Anglais | 12.1 | 14.3 | 19.0 |
| Français | 80.1 | 80.3 | 79.9 |
| Autre | 7.8 | 5.4 | 1.1 |
| Total | 100.0 | 100.0 | 100.0 |

Le groupe anglais ressort gagnant du processus: 12 % de la population a déclaré l'anglais comme langue maternelle et 14 % comme langue parlée au foyer; à l'état d'équilibre, sa part dans la population atteint 19 %. C'est évidemment l'inverse pour les tierces langues qui ne représentent guère plus d'un pourcent à l'état d'équilibre. La situation du groupe français est plus complexe. Elle met en évidence un cas de figure qui ne peut se produire quand seulement deux langues sont en présence. À s'en tenir aux compositons observées, la langue française est légèrement avantagée par le processus puisqu'elle est plus souvent parlée au foyer (80.3 %) qu'elle n'est déclarée comme langue maternelle (80.1 %); les gains qu'elle obtient aux dépens des tierces langues dépassent légèrement les pertes qu'elle fait au profit de l'anglais. Pourtant à l'état d'équilibre, la part du français (79.9 %) est légèrement inférieure à celle qu'il a comme langue maternelle. Cela tient au faible poids qu'y tiennent les tierces langues.

Si l'on disposait, en plus de la langue maternelle et de la langue d'usage, de la langue maternelle de la mère (ou du père), il serait possible de calculer un nouvel état d'équilibre où les trois compositions linguistiques seraient confondues. Peut-on considérer l'origine ethnique comme une approximation satisfaisante cette variable? Faisons cette

hypothèse. Dans cette nouvelle composition d'équilibre, la part de l'anglais augmente encore (à 19.7 %), de même que celle du français (80.3 %), tandis que celles des tierces langues devient négligeable. Cela n'est pas étranger au fait que, dans la population d'origine britannique ou française, les taux de mobilité de l'anglais ou du français vers les tierces langues sont quasi nuls. Au total, le groupe français préserve tout juste sa position relative et tout se passe comme si tout le tiers groupe était destiné à adopter l'anglais comme langue principale.

On peut donner une autre interprétation, plus classique, à la composition d'équilibre. Elle repose sur un modèle implicite de l'évolution démographique des groupes linguistiques. Supposons une population fermée aux échanges migratoires dans laquelle il n'y a ni différence de mortalité ni différence de fécondité. Admettons en outre que les taux de mobilité linguistique s'appliquent aux femmes en âge de procréer et faisons l'hypothèse que la langue maternelle des enfants correspond à la langue d'usage de leur mère. Supposons enfin que le processus est stationnaire, c'est-à-dire que les taux de mobilité ne varient pas dans le temps. Dans ce cas, la composition par langue maternelle des enfants (première génération) sera égale à la composition par langue d'usage de leur mère (génération initiale). Si l'on applique ensuite les mêmes taux de mobilité à la première génération, on obtiendra évidemment leur composition par langue d'usage; celle-ci correspondra à la composition par langue maternelle de la deuxième génération. Et ainsi de suite. Ce processus permet de suivre au fil des générations les effets des taux de mobilité sur l'évolution des compositions linguistiques. Il tend vers une composition limite qui est indépendante de l'état initial; celle-ci dépend uniquement des taux de mobilité.

La composition limite se confond évidemment avec la composition d'équilibre. On a néanmoins l'impression que le modèle sur lequel repose la composition d'équilibre est plus réaliste que celui qui engendre la composition limite. C'est que l'allure prévisionnelle de celle-ci suscite le doute. Pour éviter cette confusion, il vaut mieux s'en tenir à la notion de composition d'équilibre.

Mais il est parfois utile de situer un processus dans le temps. Cherchons à déterminer la composition par langue maternelle de la première génération (génération 1) en supposant que la fécondité (D) dans la génération initiale (génération 0) varie en fonction de la langue d'usage. Il vient:

$$n_{i.}^{(1)} = \sum_j n_{ji}^{(0)} D_i = D_i \, n_{i}^{(0)}.$$

$$\text{et} \quad n_{..}^{(1)} = D \, n_{..}^{(0)} \qquad \text{(XXXV)}$$

La fécondité moyenne (D) peut évidemment être exprimée en fonction de la fécondité propre à chaque groupe linguistique:

$$D = \sum_i p_{.i}^{(0)} D_i \qquad \text{(XXXVI)}$$

La composition par langue maternelle de la première génération dépend tout à la fois de la mobilité linguistique et des différences de fécondité:

$$p_{i.}^{(1)} = p_{.i}^{(0)} \, D_i/D \qquad \text{(XXXVII)}$$

Le processus est à l'état d'équilibre quand la composition par langue maternelle ne varie pas d'une génération à l'autre. Il suffit pour ce faire que:

$$p_{i.}^{(0)} = p_{.i}^{(0)} \, D_i/D \qquad \text{(XXXVIII)}$$

Cette condition revient à supposer que le rapport de la fécondité moyenne (D) à celle du groupe linguistique considéré ($D_i$) correspond à l'indice de continuité de ce groupe ($c_i$). Cela permet d'interpréter les effets de la mobilité linguistique en termes de fécondité différentielle (Lachapelle et Henripin, 1980).

Selon les conditions de mortalité qui prévalent actuellement au Canada, le remplacement des générations est assuré si les femmes ont en moyenne environ 2.1 enfants. Posons donc D = 2.1 et, à titre purement illustratif, calculons (à partir du tableau 3) les descendances finales qui équilibreraient les effets de la mobilité linguistique:

| | Canada | Québec | Canada moins Québec |
|---|---|---|---|
| Anglais | 1.9 | 1.8 | 1.9 |
| Français | 2.2 | 2.1 | 2.9 |
| Autre | 3.8 | 2.9 | 4.0 |
| Total | 2.1 | 2.1 | 2.1 |

Le groupe anglophone peut évidemment se contenter d'une fécondité inférieure au seuil de remplacement des générations, du fait qu'il profite des transferts linguistiques. Pour les francophones du Québec, tout se passe comme s'il n'y avait pas de mobilité linguistique: leur fécondité doit être égale à celle qui assure le remplacement des générations. Quant aux francophones et aux allophones vivant à l'extérieur du Québec, ils doivent avoir, en l'absence de migration, des descendances finales de l'ordre de 3 ou 4 enfants par femmes pour tout juste assurer le remplacement des générations. Il s'agit là évidemment d'ordres de grandeur(4).

**Vue d'ensemble**

La phase initiale de l'analyse d'un phénomène est souvent caractérisée par une grande diversité de point de vue et par l'apparition de nombreux indices. Cela permet d'en explorer les différentes facettes et de mettre en évidence les principales tendances. Mais surgissent parfois des résultats en apparence contradictoires, ce qui crée une impression de confusion et jette un doute sur l'ensemble des conclusions. C'est ainsi qu'au Québec le bilan des transferts linguistiques entre le groupe anglais et le groupe français avantage le premier au détriment du second, ce qui ne surprend guère; par contre, le taux de mobilité linguistique du groupe anglais dépasse toujours nettement celui du groupe français. Il y a là un paradoxe, du moins à première vue. Des transferts nets ou des taux de mobilité, qu'est-ce qui reflète le mieux la situation linguistique?

En l'absence d'étude méthodologique approfondie, le choix risque fort de se confondre avec une préférence subjective. Pour dépasser ce stade, il convient de procéder à un examen critique des stratégies d'analyse et des indices utilisés. Ce n'est pas là une tâche aisée. Il faut d'abord aller à l'essentiel et centrer l'attention sur ce que l'on peut appeler le tableau fondamental. S'y trouve consigné, pour une population donnée, l'effectif des individus selon la langue maternelle et la langue parlée au foyer.

Quand on s'en tient au cas le plus simple où deux langues seulement sont en présence, ce tableau comporte quatre degrés de liberté. Si l'on réduit les effectifs de manière à ce que le grand total soit égal à l'unité, on dispose de trois degrés de libertés. C'est donc dire qu'il ne peut y avoir plus de trois coefficients indépendants. Comment les choisir? On a proposé deux approches.

Les marges du tableau fondamental comportent au plus deux degrés de liberté. On épuise l'information qu'elles contiennent quand on connaît la part du groupe dominant dans la composition par langue maternelle

ainsi qu'un coefficient exprimant le déséquilibre structurel engendré par la mobilité linguistique. Plusieurs indices peuvent mettre en évidence ce déséquilibre. On a choisi le taux net de mobilité linguistique du groupe perdant. Quand ces deux coefficients sont fixés, il ne reste qu'un degré de liberté pour le noyau du tableau. Il est possible de définir un paramètre qui exprime l'intensité de la mobilité linguistique équilibrée dans la population. Tous les autres taux ou indices peuvent évidemment être exprimées en fonction des trois coefficients choisis. Cette approche se généralise facilement à un nombre quelconque de langues. Elle permet notamment de montrer qu'il n'y a rien de paradoxal dans la situation linguistique observée au Québec.

La seconde approche consiste plutôt à comparer les deux taux de mobilité aux compositions linguistiques observées. On démontre facilement que toute l'information que recèle le tableau est contenue dans les deux taux de mobilité ainsi que dans la fraction que représente l'un des groupes dans la composition par langue maternelle. Mais comment est-il possible de comparer des taux à des compositions? On a établi que les taux de mobilité peuvent être transformés en ce que l'on a appellé une composition d'équilibre. Cette transformation n'a rien d'inédit: elle correspond à un résultat classique de la théorie des chaînes de Markov stationnaires.

La première approche introduit une distinction fort pertinente entre d'une part l'effet structurel de la mobilité linguistique et d'autre part la mobilité d'échange. On s'est surtout intéressé au premier aspect au cours de la dernière décennie. Il pourrait être fructueux de considérer les deux dimensions du phénomène. Bien que la première approche semble mieux adaptée à l'étude de la mobilité linguistique que la seconde, celle-ci peut être fort utile pour aborder à des questions spéciales.

**Notes:**

(1) Dans les recensements du Canada, la langue maternelle désigne non pas la première langue parlée dans la petite enfance mais plutôt la première langue apprise (ou parlée) et encore comprise.

(2) On utilisera indifféremment dans la suite les expressions langue d'usage ou langue au foyer pour désigner la langue habituellement parlée à la maison.

(3) On montre facilement que l'on peut assimiler cette situation au cas où le groupe dominant est majoritaire. Les taux de mobilité ont en effet les mêmes intervalles de variation.

(4) Cette interprétation de l'indice de continuité a fait l'objet d'un long développement dans Lachapelle et Henripin (1980).

**Références bibliographiques**

BERNARD, Jean-Thomas, 1978. "La mobilité linguistique et l'attraction des langues au Québec", Québec, Université Laval, département d'économique (multigr.).

BOUDON, Raymond, 1973. Mathematical Structures of Social Mobility, San Francisco, Jossey-Bass, 168 p.

CASTONGUAY, Charles, 1980. "Sur quelques indices de propension à l'exogamie et au transfert linguistique", Cahiers québécois de démographie, 9, 3, 53-70.

CASTONGUAY, Charles, 1983a. "Exogamie et transferts linguistiques chez les populations de langue maternelle française au Canada", dans Démographie et destin des sous-populations, Colloque de Liège (21-23 septembre 1981), Paris, Association internatinale des démographes de langue française, 209-215.

CASTONGUAY, Charles, 1983b. Évolution des transferts linguistiques au Québec selon les recensements de 1971 et de 1981, Québec, Conseil de la langue française (édition provisoire), 76 p. (multigr.).

CASTONGUAY, Charles et MARION, Jacques, 1974. "L'anglicisation au Canada", Bulletin de l'Association des démographes du Québec, 3,1, 19-40.

DE VRIES, John et VALLEE, Frank G., 1980. Usage de la langue au Canada, Ottawa, Ministre des Approvisionnements et Services, 187 p.

GORDON, Patrick, 1965. Théorie des chaînes de Markov finies et ses applications, Paris, Dunod, 132 p.

GRYZ, Zbigniew Jan, 1980. Socio-Demographic Determinants of Language Shifts in Canada, Ph. D. Thesis, Ottawa, Carleton University, Dept. of Sociology and Anthropology, xxxi + 411 pages.

KEMENY, J.G., SNELL, J.L. et THOMPSON, G.L., 1965. Algèbre moderne et activités humaines, Paris, Dunod, 344 p.

KRALT, John, 1977. "The Case Against Net-Language Transfer: An Alternate Measure", in LAMY, Paul, Language Maintenance and Language Shift in Canada, Ottawa, University of Ottawa Press, 75-79.

LACHAPELLE, Réjean, 1980. "Evolution of Ethnic and Linguistic Composition", in BRETON, R., REITZ, J.G., and VALENTINE, V.F., Cultural Boundaries and the Cohesion of Canada, Montréal, The Institute for Research on Public Policy, 15-43.

LACHAPELLE, Réjean, 1983. "Définition et analyse des mobilités démographiques: l'exemple de la mobilité linguistique", dans Démographie et destin des sous-populations, Colloque de Liège (21-23 septembre 1981), Paris, Association internationale des démographes de langue française, 237-248.

LACHAPELLE, Réjean et HENRIPIN, Jacques, 1980. La situation démolinguistique au Canada: évolution passée et prospective, Montréal, L'Institut de recherches politiques, xxxii + 391 p.

ROCHON-LESAGE, Madeleine et MAHEU, Robert, 1974. "Composition ethnique et linguistique de la population du Québec", Annuaire du Québec, Québec, Éditeur officiel du Québec, 206-212.

## DIVISION DES ÉTUDES SOCIALES ET ÉCONOMIQUES

## SERIES D'ÉTUDES DE RECHERCHE

No.

1a. Transforming Monthly Uniform Crime Reports to Incident Records
Craig McKie, revision, Summer 1983.

1b. Transformation des données du programme de déclaration uniforme de la criminalité en fonction des données sur les cas
Craig McKie, révision, été 1983

2. The Use of "Fixed Coefficient" Matrices to Relate Occupation and Field of Study to Manpower Requirement Models
Garnett Picot, Spring 1982

3. Inconsistencies in the Occupational Coding of the Same Individual in Two Surveys
John Turner, Spring 1983

4. Sizing Up the Burden of Unemployment in Canada
Part I - A General Labour Market Perspective
R. Paul Shaw, Winter 1984

Une autre version du présent exposé est déjà parue dans Analyse de politiques, vol. XI, no. 2, juin 1985, 143-160

5. Sizing Up the Burden of Unemployment
Part II - A Perspective on Financial Hardship
R. Paul Shaw, Summer 1984

6. La ruralité québécoise: Evolutions et perspectives
Clermont Dugas, été 1984

Une autre version du présent exposé est déjà parue dans Cahiers de géographie du Québec, vol. 28, nos 73-74, avril-septembre 1984, 183-204.

7. Fiscal vs Traditional Market Variables in Intermetropolitan Migration in Canada 1956-1981
R. Paul Shaw, Summer 1984

8. Historical Statistics on Justice
Paul Reed, Summer 1984

9a. Life Table Methodology
Canada and Provinces (1980-1982 cycle)
Dhruva Nagnur, Fall 1984

9b. Méthode de calcul des tables de mortalité
Canada et provinces (cycle 1980-1982)
Dhruva Nagnur, automne 1984

10a. Heterogeneity and Selection in Population Analysis
Nathan Keyfitz, Fall 1984

10b. Hétérogénéité et effet de sélection dans l'analyse des populations
Nathan Keyfitz, automne 1984

11. The Industrial Mix of Employment: How and Why it Changes
A Review of the Past Three Decades
W. Garnett Picot, Spring 1985

12a. Analyse de la mobilité linguistique: indices, observations et modèles
Réjean Lachapelle, printemps 1985

12b. Analysis of Language Mobility: Indexes, Observations and Models
Réjean Lachapelle, Spring 1985

13a. An Analysis of Expenditure Patterns of the Elderly: Cohorts - Going Through Life Together
Louise Heslop, Spring 1985

13b. Analyse des dépenses des personnes âgées: cohorte, âge et période
Louise Heslop, printemps 1985

14a. Note de recherche - La composition linguistique: réalité et perception
Réjean Lachapelle, printemps 1985

14b. Research Note - Linguistic Composition: Perception and Reality
Réjean Lachapelle, Spring 1985

15a. Rectangularization of the Survival Curve and Entropy
The Canadian Experience
Dhruva Nagnur, Summer 1985

16a. Sizing up the Burden of Unemployment in Canada
Part III: Impact of Unemployment on Family Income
Paul Shaw, Summer 1985